Chapter 25
Weighted Least Squares for Adjusting Efficacy Data with Inconsistent Spread (78 Patients)

General Purpose

Linear regression assumes that the spread of the outcome-values is homoscedastic: it is the same for each predictor value. This assumption is, however, not warranted in many real life situations. This chapter is to assess the advantages of *weighted* least squares (WLS) instead of *ordinary* least squares (OLS) linear regression analysis.

Specific Scientific Question

The effect of prednisone on peak expiratory flow was assumed to be more variable with increasing dosages. Can it, therefore, be measured with more precision if linear regression is replaced with weighted least squares procedure.

Var 1	Var 2	Var 3	Var 4
1	29	1,40	174
2	15	2,00	113
3	38	0,00	281
4	26	1,00	127
5	47	1,00	267
6	28	0,20	172
7	20	2,00	118
8	47	0,40	383
9	39	0,40	97
10	43	1,60	304

(continued)

Var 1	Var 2	Var 3	Var 4
11	16	0,40	85
12	35	1,80	182
13	47	2,00	140
14	35	2,00	64
15	38	0,20	153
16	40	0,40	216

Var 1 Patient no
Var 2 prednisone (mg/24 h)
Var 3 peak flow (ml/min)
Var 4 beta agonist (mg/24 h)

Only the first 16 patients are given, the entire data file is entitled "weightedleastsquares" and is in extras.springer.com. SPSS 19.0 is used for data analysis. We will first make a graph of prednisone dosages and peak expiratory flows. Start with opening the data file.

Weighted Least Squares

Command:

click Graphs….Legacy Dialogs….Scatter/Dot….click Simple Scatter….click Define….Y Axis enter peakflow….X Axis enter prednisone….click OK.
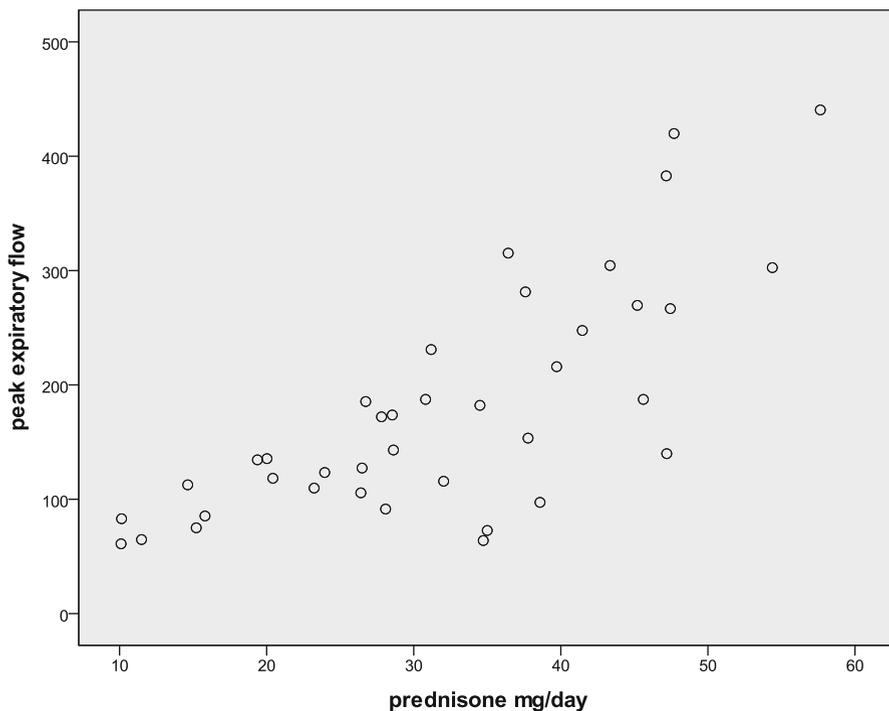
The output sheet shows that the spread of the y-values is small with low dosages and gradually increases. We will, therefore, perform both a traditional and a weighted least squares analysis of these data.

Command:

Analyze….Regression….Linear….Dependent: enter peakflow….
Independent: enter prednisone, betaagonist….OK.

Model Summary[a]				
Model	R	R square	Adjusted R square	Std. Error of the estimate
1	,763[b]	,582	,571	65,304

[a]Dependent variable: peak expiratory flow
[b]Predictors: (Constant), beta agonist mg/24 h, prednisone mg/day

Coefficients[a]						
		Unstandardized coefficients		Standardized coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	−22,534	22,235		−1,013	,314
	Prednisone mg/day	6,174	,604	,763	10,217	,000
	Beta agonist mg/24 h	6,744	11,299	,045	,597	,552

[a]Dependent variable: peak expiratory flow

In the output sheets an R value of 0.763 is observed, and the linear effects of prednisone dosages are a statistically significant predictor of the peak expiratory flow, but, surprisingly, the beta agonists dosages are not.
We will, subsequently, perform a WLS analysis.

Command:

Analyze….Regression….Weight Estimation…. select: Dependent: enter peakflow …. Independent(s): enter prednisone, betaagonist….select prednisone also as Weight variable….Power range: enter 0 through 5 by 0.5….click Options….select Save best weights as new variable….click Continue….click OK.

In the output sheets it is observed that the software has calculated likelihoods for different powers, and the best likelihood value is chosen for further analysis. When returning to the data file again a novel variable is added, the WGT_1 variable (the weights for the WLS analysis). The next step is to perform again a linear regression, but now with the weight variable included.

Command:

Analyze….Regression….Linear…. select: Dependent: enter   peakflow…. Independent(s) : enter prednisone, betaagonist….select the weights for the wls analysis (the GGT_1) variable as WLS Weight….click Save….select Unstandardized in Predicted Values….deselect Standardized in Residuals….click Continue….click OK.

Model Summary[a,b]				
Model	R	R Square	Adjusted R Square	Std. Error of the estimate
1	,846[c]	,716	,709	,125

[a]Dependent Variable: peak expiratory flow
[b]Weighted Least Squares Regression-Weighted by Weight for peakflow from WLS, MOD_6 PREDNISONE** -3,500
[c]Predictors: (Constant), beta agonist mg/24 h, prednisone mg/day

Coefficients[a,b]		Unstandardized coefficients		Standardized coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	5,029	7,544		,667	,507
	Prednisone mg/day	5,064	,369	,880	13,740	,000
	Beta agonist mg/24 h	10,838	3,414	203	3,174	,002

[a]Dependent Variable: peak expiratoryflow
[b]Weighted Least Squares Regression – Weighted by Weight for peakflow from WLS, M0D_6 PREDNISONE"-3,500

The output table now shows an R value of 0.846. It has risen from 0.763, and provides thus more statistical power. The above lower table shows the effects of the two medicine dosages on the peak expiratory flows. The t-values of the medicine predictors have increased from approximately 10 and 0.5 to 14 and 3.2. The p-values correspondingly fell from 0.000 and 0.552 to respectively 0.000 and 0.002. Larger prednisone dosages and larger beta agonist dosages significantly and independently increased peak expiratory flows. After adjustment for heteroscedasticity, the beta agonist became a significant independent determinant of peak flow.

Conclusion

The current paper shows that, even with a sample of only 78 patients, WLS is able to demonstrate statistically significant linear effects that had been, previously, obscured by heteroscedasticity of the y-value.

Note

More background, theoretical and mathematical information of weighted least squares modeling is given in Machine learning in medicine part three, Chap. 10, Weighted least squares, pp 107–116, Springer Heidelberg Germany, 2013, from the same authors.